



United States  
Environmental Protection  
Agency

---

# EPA's Review of Industrial Wastewater Discharge Monitoring Report (DMR) Data for Preliminary Program Plan 15

September 2021

---

THIS PAGE INTENTIONALLY LEFT BLANK

---

U.S. Environmental Protection Agency  
Office of Water (4303T)  
1200 Pennsylvania Avenue, NW  
Washington, DC 20460

*Docket EPA-HQ-OW-2021-0547*  
*[www.regulations.gov](http://www.regulations.gov)*  
*DCN 11103*

---

THIS PAGE INTENTIONALLY LEFT BLANK

## Contents

---

1. Introduction.....	1
2. Cross-Category Concentration Analysis Data Sources and Methodology .....	1
2.1 Data Supporting the Cross-Category Concentration Analysis .....	1
2.1.1 Use and Considerations of Discharge Monitoring Report Data .....	1
2.1.2 Linking Facilities to PSCs .....	3
2.1.3 2019 DMR Data Quality Review.....	3
2.2 Cross-Category Concentration Analysis Methodology .....	4
2.2.1 Data Preparation .....	5
2.2.2 Methodology.....	6
2.2.3 Cross-Category Concentration Analysis Limitations.....	9
3. Results of the Cross-Category Concentration Analysis.....	10
4. References.....	15

## List of Tables

---

Table 1. Results of the 2019 DMR Data Completeness Check .....	4
Table 2. Median Concentrations of Pollutants Reported by Facilities in the Sugar Processing PSC .....	6
Table 3. PSCs Sorted by Median Ammonia as N Concentration <sup>a</sup> .....	7
Table 4. Top Five Pollutant Rank for the Sugar Processing PSC .....	8
Table 5. Top 25 Percent Pollutant Rank for the Sugar Processing PSC.....	9
Table 6. Cross-Category Concentration Analysis Results .....	12

## List of Acronyms

---

CWA	Clean Water Act
DMR	Discharge Monitoring Report
ECHO	Enforcement and Compliance History Online
ELG	Effluent limitations guidelines and standards
GPD	Gallons per day
ICIS-NPDES	Integrated Compliance Information System for the National Pollutant Discharge Elimination System
kg/d	Kilograms per day
Loading Tool	Water Pollutant Loading Tool
MGD	Million gallons per day
mg/L	Milligram per liter
NAICS	North American Industry Classification System
NODI	No Data Indicator
NPDES	National Pollutant Discharge Elimination System
PSC	Point Source Category
SIC	Standard Industrial Classification
WET	Whole Effluent Toxicity

# 1. Introduction

---

Effluent limitations guidelines and standards (ELGs) are an essential element of the nation's clean water program, established by the 1972 amendments to the Clean Water Act (CWA). ELGs are technology-based regulations used to control pollution in industrial wastewater discharges. This regulatory program substantially reduces industrial wastewater pollution and continues to be a critical aspect of the effort to clean the nation's waters.

Per CWA requirements, EPA annually reviews all industrial point source dischargers to identify existing ELGs that are potential candidates for revision<sup>1</sup> and to identify new point source categories (PSCs) for the potential development of ELGs.<sup>2</sup> As part of EPA's 2020 review of ELGs, EPA evaluated 2019 concentration data reported by industrial facilities on discharge monitoring reports (DMRs). This analysis, referred to as the cross-category concentration analysis, compared facility wastewater discharge pollutant concentrations across industrial PSCs to identify categories that have relatively high pollutant concentration discharges compared to other PSCs and provided a means of prioritizing specific PSCs for further review and study. EPA first performed this analysis for its 2019 annual review of ELGs using 2017 concentration data. See *EPA's Review of Industrial Wastewater Discharge Monitoring Report (DMR) Data* for details of the 2019 annual review analysis (U.S. EPA, 2020). EPA used updated DMR data and generally employed the same methodology for this current review, with minor adjustments as discussed in this report.

This report presents the data quality review, methodology, considerations, findings, and next steps from EPA's cross-category concentration analysis using 2019 DMR concentration data.

## 2. Cross-Category Concentration Analysis Data Sources and Methodology

---

EPA analyzed pollutant discharge concentrations using publicly available industrial wastewater discharge data to rank categories by the concentration of pollutants in their discharges relative to other PSCs. This section discusses the data sources and the methodology used for the cross-category concentration analysis.

### 2.1 Data Supporting the Cross-Category Concentration Analysis

For this analysis, EPA evaluated available industrial wastewater discharge data reported on DMRs. Facilities that directly discharge wastewater to surface waters of the United States pursuant to a National Pollutant Discharge Elimination System (NPDES) permit are required to report monitoring data via DMRs for pollutants listed in their NPDES permits. Facilities send DMRs electronically to their respective NPDES permitting authority (state or EPA). The DMR data are stored in EPA's centralized program database, Integrated Compliance Information System for the National Pollutant Discharge Elimination System (ICIS-NPDES).

#### 2.1.1 Use and Considerations of Discharge Monitoring Report Data

ICIS-NPDES captures pollutant-specific permit limits, monitoring requirements, and DMR data, including, but not limited to, facility-, outfall-, and monitoring-period-specific pollutant discharge concentrations, quantities, and wastewater flows. With more than 375,000 industrial facilities permitted for wastewater discharges to waters of the United States in 2019 (ERG, 2021), the ICIS-NPDES database continues to be

---

<sup>1</sup> See CWA sections 304(b), 301(d), 304(m)(1)(A) and 304(g), 33 U.S.C. 1314(b), 1311(d), 1314(m)(1)(A) and 1314(g).

<sup>2</sup> See CWA sections 304(m)(1)(B), 33 U.S.C. 1314(m)(1)(B), and CWA section 307(b), 33 U.S.C. 1317(b).

the most comprehensive data source quantifying pollutants discharged directly to surface waters of the U.S.

The data collected in the ICIS-NPDES system are particularly useful for ELG planning and this analysis for the following reasons:

- ICIS-NPDES is national in scope, including data from all 50 states and 21 U.S. territories/tribes.
- DMR data included in ICIS-NPDES are based on effluent chemical analysis and metered flows using known analytical methods.
- ICIS-NPDES includes discharge data for all facilities with NPDES permits; therefore, the data are not limited to certain industries.

For these reasons, EPA has historically used DMR data for its annual reviews as a screening tool to evaluate industrial wastewater discharges. EPA identified the following limitations of DMR data collected in the ICIS-NPDES data system:

- ICIS-NPDES contains data only for pollutants that a facility is required by its permit to monitor; the facility is not required to monitor or report all pollutants discharged.
- ICIS-NPDES does not include data characterizing discharges from industrial facilities to POTWs (indirect discharges).
- Facilities enter DMR data manually, which sometimes results in transcription errors (e.g., data entered into the wrong field or reported with an incorrect unit of volume or time).
- ICIS-NPDES contains data on the permitted feature (e.g., external outfall) but does not explicitly identify the type of wastewater being discharged (e.g., process wastewater, stormwater, noncontact cooling water). In some cases, this information may be deduced from the name or description of the outfall reported by the facility; however, total flow rates reported may include non-process wastewater, such as stormwater and noncontact cooling water, as well as process wastewater.

Despite these limitations, EPA determined that the ICIS-NPDES data are a robust and reliable source of information on industrial wastewater discharges, particularly for this initial screening-level review. EPA considered these limitations while developing the methodology of the cross-category concentration analysis discussed throughout this report.

EPA downloaded the following three sets of DMR data from ICIS-NPDES for calendar year 2019:

- 2019 DMR Industrial Monthly Average Concentration Data (ERG, 2020b)
- 2019 DMR Industrial Monthly Average Quantity Data (ERG, 2020c)
- 2019 DMR Flow Data (ERG, 2020d)

EPA used 2019 data for this review because they were the most recent and complete set of industrial wastewater discharge data available when this review began.

To simplify the analysis and focus the cross-category concentration analysis on toxic and nonconventional pollutants, EPA excluded conventional pollutants,<sup>3</sup> pollutants in drilling fluid, pollutants measured in units that are not comparable with units for the concentration or quantity data (e.g., percent), and whole effluent toxicity (WET) parameters (ERG, 2020e). EPA may choose to include these pollutants in subsequent reviews. The raw data sets listed above include the pollutants reported by facilities.

---

<sup>3</sup> CWA section 304(a)(4) designates the following as conventional pollutants: biochemical oxygen demand (BOD<sub>5</sub>), total suspended solids, fecal coliform, pH, and any additional pollutants defined by the Administrator as conventional. The Administrator designated oil and grease as an additional conventional pollutant on July 30, 1979 (44 FR 44501).

Facilities may monitor and report concentration and quantity data for different statistical bases (i.e., averages, maximums, or minimums) and frequencies (e.g., annually, monthly, or daily) depending on their NPDES permit requirements. To maintain comparability between data reported by facilities and account for variability of the data throughout the year, EPA used concentration and quantity data reported as monthly averages in this analysis. See Appendix B of *EPA's Review of Industrial Wastewater Discharge Monitoring Report (DMR) Data* (EPA's cross-category concentration analysis report for the 2019 annual review) for more information on the statistical bases and methodology for identifying monthly average concentrations (U.S. EPA, 2020).

For each of the three DMR datasets downloaded from ICIS-NPDES, EPA captured facility information for the list of included pollutants for each monitoring period in 2019. Facility information included the following elements at the least; additional data were available for some facilities:

- NPDES permit number.
- Standard Industrial Classification (SIC) code.
- External outfall number.
- Permit limit set identifier,<sup>4</sup> if applicable.
- Permit limit, if applicable.
- Pollutant name.
- Monitoring period dates.
- Monthly average concentration (mg/L), monthly average quantity (kg/d), or all reported flow (MGD) values, as available, respective to the datasets listed above.
- No Data Indicator (NODI) code.<sup>5</sup>

EPA imported and processed the datasets in a static database to preserve the integrity of the data and facilitate subsequent analyses (ERG, 2020a).

### **2.1.2 Linking Facilities to PSCs**

EPA used established crosswalks maintained in the [Water Pollutant Loading Tool](#) (Loading Tool) documentation to relate individual facility and reported pollutants to the most appropriate PSC, commonly based on the facility's primary reported SIC or North American Industry Classification System (NAICS) code.<sup>6</sup> These links enable EPA to analyze discharges within and across PSCs. The Loading Tool was developed by EPA to explore and analyze DMR data. See Section 3 of the *Technical Users Background Document for the Discharge Monitoring Report (DMR) Pollutant Loading Tool* (Loading Tool Technical Users Document) for more information on these crosswalks (U.S. EPA, 2012). EPA updates and refines the facility to PSC crosswalk as it gathers updated information, or, in some cases, performs a facility-specific assessment of the process operations generating wastewater. The crosswalk is available for download on the [Loading Tool Resources](#) webpage.

### **2.1.3 2019 DMR Data Quality Review**

For this analysis, EPA evaluated the completeness, accuracy, and reasonableness of the downloaded 2019 DMR data as follows.

---

<sup>4</sup> A permit limit set is a unique identifier associated with a group of limits in a permit. This identifier is particularly useful when a permit limit includes multiple limits for the same pollutant(s).

<sup>5</sup> The no data indicator (NODI) code indicates the reason that data for an expected DMR value were not submitted by the permittee for a monitoring period end date.

<sup>6</sup> EPA did not review facilities that do not have an industrial classification (did not report a SIC code), facilities that report a SIC code of 4952 (publicly and privately-owned treatment works), and facilities that report a SIC code but are not industrial facilities.

*Completeness.* EPA assessed completeness of the datasets by comparing the volume of the 2019 downloaded ICIS-NPDES data to that of the 2017 data to ensure that there was no discrepancy that would indicate an incomplete download of the overall data. Table 1 compares the overall count of facilities and records reported in 2017 and 2019.

As expected, the 2019 concentration and quantity datasets were larger than the 2017 dataset, as new facilities or pollutants may be added each year as permits are developed or revised.

**Table 1. Results of the 2019 DMR Data Completeness Check**

Check	Data Type	2017 Count	2019 Count	Percent Change from 2017 to 2019
Total size of dataset (Number of Records)	Concentration	1,978,642	2,199,584	10% Increase
	Quantity	308,081	309,832	0.57% Increase
Number of Facilities	Concentration	30,566	31,239	2.2% Increase
	Quantity	12,412	12,323	0.72% Decrease

Sources: ERG, 2020b, 2020c, 2020d.

*Accuracy and reasonableness.* EPA also evaluated the accuracy and reasonableness of the data. EPA compared the smallest and largest reported monthly average concentrations or quantities within a monitoring period from 20 randomly selected facilities representing a variety of states and industries, with the data on EPA’s [Enforcement and Compliance History Online \(ECHO\) website](#) to ensure that the values were accurate. Subsequently, for each of these facilities, EPA compared the smallest and largest reported concentrations or quantities to monthly average concentration or quantity for a pollutant across several years of reported data using effluent charts<sup>7</sup> on the ECHO website. See *2019 DMR Data Quality Checks* for the results of this data quality check (ERG, 2020f). EPA did not identify any inconsistencies with data for these 20 facilities for 2019, nor in any other reporting years, indicating that the data downloaded are accurate and reasonable.

Because the reported flow values affect the calculated concentration for the quantity-based data (see Section 2.1.1 for the method used to calculate concentrations), EPA reviewed and corrected flow values exceeding 5,000 MGD. Consistent with the methodology developed in the Loading Tool (see Section 3.1.2 in the Loading Tool Technical Users Document), EPA assumed that flow values above 5,000 MGD were incorrectly reported in units of gallons per day. As a result, EPA corrected these values by dividing the flow by 1,000,000. EPA corrected flow values from 76 facilities within the static database and used the corrected data in its analyses (ERG, 2020a).

Based on these checks, EPA determined that the downloaded 2019 DMR data were useable for the cross-category concentration analysis. EPA will conduct a separate quality review of the data for the PSCs that are prioritized for further review.

## 2.2 Cross-Category Concentration Analysis Methodology

This section discusses the preparation of the data for use in the analysis (Section 2.2.1), the methodology for prioritizing point source categories for further review (Section 2.2.2), and considerations of the analysis (Section 2.2.3).

<sup>7</sup> A facility’s ECHO effluent charts may be accessed through a Facility Name/ID search and clicking on the “E” link under Reports in the search results (<https://echo.epa.gov/>).

### 2.2.1 Data Preparation

To prepare the data for analysis, EPA: (1) quantified data reported as non-detect, and (2) calculated discharged concentrations of pollutants from reported quantity and flow data (when reported concentration data were not available). EPA then combined all reported and calculated concentration data for use in its cross-category concentration analysis.

#### Quantifying Data Reported as Below Detection Limits

Facilities may report a monitoring period value (in this case, the monthly average concentration or quantity) as below detection in one of two ways: (1) enter the detection limit as the value, preceded by a “less than” (<) symbol or (2) report a NODI value of “B” to indicate the measured value is below the detection limit. Consistent with its approach in previous ELG planning reviews, EPA handled data reported below the detection limit as follows:

- If a facility reported all monthly average concentrations or quantities in 2019 as below the detection limit for a pollutant, EPA used a concentration or quantity of zero for that pollutant.
- If a facility reported some monthly average concentrations or quantities as below the detection limit and other values above the detection limit in 2019 for a pollutant, EPA replaced the non-detected values with a value equal to one-half the detection limit, because the pollutant is reasonably expected to be present in the facility’s wastewater.

#### Calculating Concentrations from Reported Quantity Data

Depending on the permit requirements, some facilities may report wastewater discharges as quantities discharged over time (e.g., kg/d) rather than concentrations. To directly compare all discharges within and across pollutants and PSCs, EPA calculated discharge concentrations of pollutants from reported quantity and flow data, if necessary, as described below.

EPA identified the reported flow rate value associated with each reported quantity value by matching the NPDES permit, outfall, monitoring period, and permit limit set.<sup>8</sup>

1. EPA excluded flows that did not represent industrial process waste streams, inferred from the flow name. If the flow parameter was identified as an overflow volume, stormwater flow, pump outflow, ballast water flow, and sanitary waste flow, EPA excluded the flow value from the cross-category concentration analysis.
2. Of the flows identified as industrial process wastewater streams, if the quantity value lacked a corresponding flow value, or if the flow value was zero, EPA did not calculate a concentration from the quantity data, and the value was excluded from the cross-category concentration analysis. There were 16,035 quantity records (approximately 5 percent of quantity records) that lacked a corresponding flow value or reported a flow value of zero.
3. If the quantity value had exactly one corresponding flow value, EPA used the flow value and the reported quantity value to calculate a concentration.
4. If the quantity value had more than one corresponding flow value for a monitoring period (e.g., the facility reported an average, total, and/or maximum flow), EPA selected flow values based on a hierarchy of the statistical bases (e.g., average, total, maximum) and flow parameter description. This hierarchy, developed for the Loading Tool, prioritizes use of average flow values, then the total, then the maximum flow values, when available. See Appendix E of *EPA’s Review of Industrial Wastewater Discharge Monitoring Report (DMR) Data* (EPA’s cross-category concentration analysis report for the 2019 annual review) for the table showing the hierarchy of

---

<sup>8</sup> A permit limit set is a unique identifier associated with a group of limits in a permit. This identifier is particularly useful when a permit limit includes multiple limits for the same pollutant(s).

flow parameters used to calculate concentrations from reported quantity data. EPA used the prioritized flow value and the reported quantity to calculate a concentration.

### Combining Reported and Calculated Concentrations into a Dataset

EPA combined the calculated monthly average concentration data with reported monthly average concentration data for all facilities and all monitoring periods (considering quantified non-detect values) into a static database for use in the cross-category concentration analysis. If a facility reported both a concentration and a quantity for the same monitoring period, parameter, and outfall, EPA prioritized the reported concentration over the calculated concentration derived from the quantity value to avoid double counting data. EPA then averaged the monthly average concentrations from 2019 (both reported and calculated) to calculate a single 2019 average monthly concentration for each pollutant reported for each facility that could be compared with other facilities for use in the cross-category concentration analysis (ERG, 2020a).

### 2.2.2 Methodology

The cross-category concentration analysis consists of the following steps described in this section. This analysis compares wastewater discharge pollutant concentrations across PSCs for multiple pollutants to identify categories that have relatively higher pollutant concentration discharges compared to other PSCs and provides a means of prioritizing specific PSCs for further review and study. Following each of the steps, EPA presents an example of the cross-category analysis methodology as applied to one pollutant or PSC, depending on the step.

#### Step 1: Calculate Median Pollutant Concentrations by PSC

From the concentration dataset, EPA grouped facilities into PSCs using the facility-to-PSC crosswalk (as described in Section 2.1.2) and calculated the median of the average monthly concentrations (hereafter referred to as the median concentration) for each pollutant discharged by facilities in each PSC. EPA narrowed the scope of this analysis using the following assumptions:

- EPA excluded pollutants that were only reported by one facility within a PSC from this analysis because they are unlikely to be representative of discharges within a PSC.
- EPA only ranked PSCs where median concentrations were greater than 0 mg/L to focus its review on top-ranking discharges.
- EPA removed pollutants reported by only one PSC because the analysis is intended to provide a comparison of discharge concentrations across PSCs.

As an example, EPA presents the cross-category concentration analysis for pollutant discharges associated with Sugar Processing (40 CFR Part 409). This PSC has discharge concentration data for 34 pollutants; 22 are reported by only one facility and are excluded from further analysis. Table 2 presents the median concentration for all the pollutants included in the analysis for the Sugar Processing PSC.

**Table 2. Median Concentrations of Pollutants Reported by Facilities in the Sugar Processing PSC**

Pollutant <sup>a</sup>	Median Concentration <sup>b</sup>	Units of Median Concentration	Number of Facilities Reporting Pollutant
Ammonia as N	3.62	mg/L	11
Bicarbonate ion (as HCO <sub>3</sub> )	498	mg/L	2
Chloride	88.0	mg/L	2
Copper	23.9	µg/L	2
Lead	4.46	µg/L	2
Magnesium	42.6	mg/L	2
Mercury	0.138	µg/L	2

**Table 2. Median Concentrations of Pollutants Reported by Facilities in the Sugar Processing PSC**

Pollutant <sup>a</sup>	Median Concentration <sup>b</sup>	Units of Median Concentration	Number of Facilities Reporting Pollutant
Nitrogen	12.4	mg/L	3
Phosphorus	0.591	mg/L	12
Sodium	159	mg/L	2
Total Kjeldahl Nitrogen	12.3	mg/L	2
Total Residual Chlorine	42.2	µg/L	2

a – This list excludes the following pollutants: aldicarb, arsenic, base neutrals and acid, boron, cadmium, chlorodibromomethane, chloropyrifos, cyanide, di(2-ethylhexyl) phthalate, total hardness, inorganic nitrogen, iron, manganese, nickel, potassium, selenium, silver, sulfate, undissociated hydrogen sulfide, 2,3,7,8-tetrachlorodibenzodioxin equivalents, volatile organic compounds, and zinc.

b – EPA rounded the median concentration to three significant figures.

### Step 2: Identify PSCs with Highest Median Concentrations by Pollutant

For each of the pollutants included in the analysis from Step 1, EPA sorted the median pollutant concentrations for the PSCs from highest to lowest and assigned the PSC a rank.

For example, 51 PSCs reported ammonia as N. Table 3 presents the top 15 PSCs sorted from highest to lowest median concentration for ammonia as N and the corresponding EPA-assigned rank.

**Table 3. PSCs Sorted by Median Ammonia as N Concentration<sup>a</sup>**

PSC Rank by Median Concentration of Ammonia as N	PSC Name	Median Ammonia as N Concentration (mg/L) <sup>b</sup>
1	Explosives Manufacturing	13.0
2	Centralized Waste Treatment	6.71
3	Electrical and Electronic Components	5.93
4	Canned and Preserved Seafood Processing	4.61
5	Sugar Processing	3.62
6	Waste Combustors	2.91
7	Fertilizer Manufacturing	1.96
8	Soap And Detergent Manufacturing	1.25
9	Oil & Gas Extraction	0.969
10	Grain Mills	0.923
11	Inorganic Chemicals Manufacturing	0.880
12	Pulp, Paper and Paperboard	0.856
13	Unassigned Waste Facility	0.841
14	Textile Mills	0.823
15	Landfills	0.788

a – This is not the complete list of the PSCs reporting ammonia as N.

b – EPA rounded the median concentration to three significant figures.

### Step 3: Calculate PSC Scores

For the 2020 cross-category concentration analysis, EPA used two approaches to assess the relative number of top-ranking pollutants within a PSC and from the results, developed a PSC score for each approach.

- *Top Five PSC Approach:* Counts the number and percent of pollutants where the median concentration for the PSC was among the five highest median concentrations for the pollutant across all PSCs (see Step 3a below).
- *Top 25 Percent PSC Approach:* Counts the number and percent of pollutants where the median concentration for the PSC was among the top 25 percent of highest median concentrations for the pollutant across all PSCs (see Step 3b below).

In its prior 2019 cross-category concentration analysis (U.S. EPA, 2020), EPA ranked PSCs using only the *Top Five PSC Approach*. Including the *Top 25 Percent Approach* in the 2020 rankings analysis allowed EPA to consider and better account for PSCs with unique pollutants that are reported by very few categories, which may be overweighted in the *Top Five PSC Approach*.

To normalize for the varying number of pollutants reported by each PSC, for each approach EPA divided the count of top-ranking pollutants by the total number of pollutants reported by more than one facility in the PSC. This provided a directly comparable “score” for each PSC representing the percent of pollutants in the PSC with median concentrations ranked higher across PSCs.

#### *Step 3a: Calculating a PSC Score Using the Top Five PSC Approach*

For each PSC, EPA counted the number of pollutants where the median concentration for the PSC ranked among the five highest median concentrations for the pollutant across all PSCs (from Step 2). EPA then divided the count of top-ranking pollutants by the total number of pollutants reported by more than one facility in the PSC to calculate the overall score (the percent of top-ranking pollutants) for the PSC.

Using the Sugar Processing Category to illustrate this approach, the median pollutant concentrations ranked among the top five across PSCs for seven of the 12 pollutants (58 percent) reported by the facilities in the PSC, as shown in Table 4. This percentage becomes the PSC score for this approach.

**Table 4. Top Five Pollutant Rank for the Sugar Processing PSC**

Pollutant Name	Median Concentration <sup>a</sup>	Median Sugar Processing Concentration Rank Compared to Other PSCs
Mercury	0.138 µg/L	1 of 28
Bicarbonate ion (as HCO <sub>3</sub> )	498 mg/L	2 of 5
Sodium	159 mg/L	3 of 10
Total Kjeldahl Nitrogen	12.3 mg/L	4 of 36
Ammonia as N	3.62 mg/L	5 of 51
Copper	23.9 µg/L	5 of 44
Lead	4.46 µg/L	5 of 30
Magnesium	42.6 mg/L	6 of 11
Total Residual Chlorine	0.0422 mg/L	7 of 32
Nitrogen	12.4 mg/L	11 of 36
Phosphorus	0.591 mg/L	16 of 47
Chloride	88.0 mg/L	21 of 35

a – EPA rounded the median concentration to three significant figures.

#### *Step 3b: Calculating a PSC Score Using the Top 25 Percent PSC Approach*

For each pollutant, EPA identified PSCs with median concentrations that ranked among the top 25 percent for the pollutant across PSCs, using the following equation:

$$\text{Top 25 Percent} = \text{Round Up} [\text{Total Number of PSCs Reporting Pollutant} \times 0.25]^9$$

For example, 51 PSCs reported ammonia as N; therefore, EPA counted the top 13 PSCs with the highest median ammonia as N concentration.

For each PSC, EPA counted the number of pollutants where the PSC ranked among the top 25 percent of median concentrations across PSCs. EPA then divided the count of top-ranking pollutants by the total number of pollutants reported by more than one facility in the PSC to calculate the overall score (the percent of top-ranking pollutants) for the PSC.

Continuing to use the Sugar Processing Category to illustrate this approach, the median pollutant concentrations ranked among the top 25 percent of PSCs for eight of the 12 pollutants (67 percent) reported by the facilities in the PSC, as shown in Table 5. This percentage becomes the PSC's score.

**Table 5. Top 25 Percent Pollutant Rank for the Sugar Processing PSC**

Pollutant Name	Number of PSCs Comprising Top 25 Percent of Median Concentrations <sup>a</sup>	Median Sugar Processing Concentration Rank Compared to Other PSCs
Mercury	7	1 of 28
Bicarbonate ion (as HCO <sub>3</sub> )	2	2 of 5
Sodium	3	3 of 10
Total Kjeldahl Nitrogen	9	4 of 36
Ammonia as N	13	5 of 51
Copper	11	5 of 44
Lead	8	5 of 30
Total Residual Chlorine	8	7 of 32
Magnesium	3	6 of 11
Nitrogen	9	11 of 36
Phosphorus	12	16 of 47
Chloride	9	21 of 35

a – EPA used the following formula to calculate this value: Top 25 Percent = Round Up[Total Number of PSCs Reporting Pollutant x 0.25].

#### Step 4: Rank and Prioritize PSCs for Further Review

EPA ranked the PSCs by the scores generated from both the *Top Five PSC Approach* and the *Top 25 Percent PSC Approach* (identified in Step 3). EPA selected the top five ranking PSCs from each approach for further consideration for a preliminary category review, excluding any PSCs currently being reviewed as identified in Effluent Guidelines Program Plan 14. Section 3 (Table 6) presents the results of the cross-category concentration analysis for both approaches.

#### 2.2.3 Cross-Category Concentration Analysis Limitations

EPA identified several limitations of the cross-category concentration analysis, which include, but are not limited to the following.

- Analysis is relative to what other categories are reporting and does not consider the extent of discharge. A PSC that discharges larger concentrations relative to other categories may or may not indicate the potential for reducing or eliminating pollutant discharges within that PSC.

<sup>9</sup> EPA rounded up to the next whole number to ensure a minimum of one PSC was considered for each pollutant.

- Analysis uses median concentration and does not directly account for the range of concentration data within a PSC.
- Analysis does not compare the median pollutant concentrations for a PSC to any national effluent limitations, if there is one, or to specific permit limits.
- Analysis does not consider the magnitude (i.e., pollutant loading) or toxicity of the pollutants being discharged.
- Analysis may rank higher those PSCs whose facilities monitor and report pollutants unique to the PSC simply because few other PSCs report those pollutants; though also considering categories that rank high using the *Top 25 Percent PSC Approach* addresses this limitation to some extent.

Even with these limitations, EPA considered the cross-category concentration analysis an appropriate method to provide a screening-level review of industrial discharges, as it provided an indication of the extent to which a PSC has larger concentrations of pollutant discharges relative to other PSCs. This analysis considered all DMR data reported as concentration and quantity simultaneously, including facilities with monitoring requirements only. The cross-category concentration analysis was a starting point for prioritizing PSCs for further review. To the extent possible, EPA addressed the limitations associated with the analysis as part of the review of prioritized PSCs, which included a review of the range and magnitude of concentrations and comparison to national standards and treatment performance data.

### 3. Results of the Cross-Category Concentration Analysis

---

Table 6 presents the results of this analysis, including the following information for each PSC. Though the table presents results from the *Top Five PSC Approach* and *Top 25 Percent PSC Approach*, it is sorted from highest to lowest score from the *Top Five PSC Approach* (and then alphabetically for PSCs with the same score).

- *Top Five PSC Approach PSC Score (Column A)*. Percent of pollutants reported by more than one facility where the PSC's median concentration ranked among the top five highest median concentrations reported for the pollutant across all PSCs. Value is calculated from the number of pollutants that rank in the top five (Column B) and the number of pollutants with data reported (Column E).
- *Top Five PSC Approach Number of Top-Ranking Pollutants (Column B)*. Number of pollutants reported by more than one facility where the PSC's median concentration ranked among the top five highest median concentrations reported for the pollutant across all PSCs.
- *Top 25 Percent PSC Approach PSC Score (Column C)*. Percent of pollutants reported by more than one facility where the PSC's median concentration ranked among the top 25 percent highest median concentrations reported for the pollutant across all PSCs. Value is calculated from the number of pollutants that rank in the top 25 percent (Column D) and the number of pollutants with data reported (Column E).
- *Top 25 Percent PSC Approach Top-Ranking Pollutants (Column D)*. Number of pollutants reported by more than one facility where the PSC's median concentration ranked among the top 25 percent highest median concentrations reported for the pollutant across all PSCs.
- *Number of Pollutants with Data Reported (Column E)*. Number of pollutants that were reported by more than one facility within a PSC and a discharge concentration of greater than zero, and therefore, considered in the cross-category concentration analysis for the PSC.
- *Number of Facilities Reporting Data (Column F)*. Total number of facilities reporting data for any pollutant within the PSC (including pollutants that were reported by only one facility).

EPA excluded from further consideration categories currently under review (U.S. EPA, 2021). For the remaining categories, EPA identified the five PSCs with the highest scores from both the *Top Five PSC Approach* (highlighted in blue in Table 6) and *Top 25 Percent PSC Approach* (highlighted in green in Table 6) for further review. Three of the top PSCs overlapped between the two analyses, therefore EPA selected for further review seven total PSCs, which included:

- Canned and Preserved Seafood Processing (40 CFR Part 408)
- Sugar Processing (40 CFR Part 409)
- Soap and Detergent Manufacturing (40 CFR Part 417)
- Metal Products and Machinery (40 CFR Part 438)
- Paint Formulating (40 CFR Part 446)
- Explosives Manufacturing (40 CFR Part 457)
- Battery Manufacturing (40 CFR Part 461)

**Table 6. Cross-Category Concentration Analysis Results**

40 CFR Part	PSC Name	Top Five PSC Approach		Top 25 Percent PSC Approach		Number of Pollutants with Data Reported	Number of Facilities Reporting Data
		PSC Score (Percent of Top- Ranking Pollutants)	Number of Top-Ranking Pollutants	PSC Score (Percent of Top-Ranking Pollutants)	Number of Top-Ranking Pollutants		
		A	B	C	D		
469	Electrical and Electronic Components <sup>a</sup>	100.0%	3	100.0%	3	3	5
461	Battery Manufacturing	100.0%	1	100.0%	1	1	2
438	Metal Products and Machinery	100.0%	11	0.0%	0	11	63
457	Explosives Manufacturing	80.0%	4	80.0%	4	5	6
408	Canned and Preserved Seafood Processing	66.7%	6	77.8%	7	9	27
417	Soap and Detergent Manufacturing	63.3%	19	20.0%	6	30	11
429	Timber Products Processing	60.5%	23	28.9%	11	38	54
409	Sugar Processing	58.3%	7	66.7%	8	12	15
455	Pesticide Chemicals	55.6%	5	55.6%	5	9	15
414	Organic Chemicals, Plastics and Synthetic Fibers	51.6%	32	24.2%	15	62	296
437	Centralized Waste Treatment <sup>a</sup>	50.0%	14	32.1%	9	28	7
446	Paint Formulating	50.0%	2	75.0%	3	4	6
443	Paving and Roofing Materials (Tars and Asphalt)	50.0%	8	50.0%	8	16	37
420	Iron and Steel Manufacturing	44.7%	21	38.3%	18	47	100
432	Meat and Poultry Products <sup>a</sup>	43.3%	13	40.0%	12	30	185
415	Inorganic Chemicals Manufacturing	40.5%	17	38.1%	16	42	112
467	Aluminum Forming	40.0%	4	40.0%	4	10	10
N/A	Food Service Establishments	40.0%	2	40.0%	2	5	107
433	Metal Finishing	40.0%	18	26.7%	12	45	357
439	Pharmaceutical Manufacturing	37.5%	9	37.5%	9	24	32
430	Pulp, Paper and Paperboard	37.1%	13	40.0%	14	35	145
421	Nonferrous Metals Manufacturing	37.0%	10	40.7%	11	27	36
426	Glass Manufacturing	36.8%	7	42.1%	8	19	23
442	Transportation Equipment Cleaning	36.4%	8	18.2%	4	22	39

**Table 6. Cross-Category Concentration Analysis Results**

40 CFR Part	PSC Name	Top Five PSC Approach		Top 25 Percent PSC Approach		Number of Pollutants with Data Reported	Number of Facilities Reporting Data
		PSC Score (Percent of Top- Ranking Pollutants)	Number of Top-Ranking Pollutants	PSC Score (Percent of Top-Ranking Pollutants)	Number of Top-Ranking Pollutants		
		A	B	C	D		
423	Steam Electric Power Generating <sup>a</sup>	36.0%	18	10.0%	5	50	442
445	Landfills	34.1%	14	19.5%	8	41	143
N/A	Independent and Stand Alone Labs	33.3%	4	25.0%	3	12	14
435	Oil and Gas Extraction <sup>a</sup>	33.3%	9	29.6%	8	27	76
449	Airport Deicing	31.3%	5	37.5%	6	16	44
436	Mineral Mining and Processing	31.3%	10	18.8%	6	32	217
471	Nonferrous Metals Forming and Metal Powders	30.4%	7	56.5%	13	23	33
405	Dairy Products Processing	30.0%	6	30.0%	6	20	77
418	Fertilizer Manufacturing	27.8%	5	33.3%	6	18	35
N/A	Drinking Water Treatment	27.8%	10	22.2%	8	36	1425
N/A	Unassigned Waste Facility	27.5%	11	17.5%	7	40	115
460	Hospital	26.7%	4	33.3%	5	15	140
468	Copper Forming	25.0%	1	50.0%	2	4	5
424	Ferroalloy Manufacturing	23.8%	5	23.8%	5	21	9
440	Ore Mining and Dressing	23.3%	7	10.0%	3	30	72
419	Petroleum Refining	21.9%	7	15.6%	5	32	331
434	Coal Mining	20.5%	8	10.3%	4	39	1700
412	Concentrated Animal Feed Operations	20.0%	1	40.0%	2	5	16
450	Construction and Development	20.0%	4	25.0%	5	20	49
464	Metal Molding and Casting (Foundries)	20.0%	3	40.0%	6	15	29
407	Canned and Preserved Fruits and Vegetables Processing	18.8%	3	25.0%	4	16	56
444	Waste Combustors	13.3%	2	33.3%	5	15	11
451	Concentrated Aquatic Animal Production	12.5%	2	0.0%	0	16	193

**Table 6. Cross-Category Concentration Analysis Results**

40 CFR Part	PSC Name	Top Five PSC Approach		Top 25 Percent PSC Approach		Number of Pollutants with Data Reported	Number of Facilities Reporting Data
		PSC Score (Percent of Top- Ranking Pollutants)	Number of Top-Ranking Pollutants	PSC Score (Percent of Top-Ranking Pollutants)	Number of Top-Ranking Pollutants		
		A	B	C	D		
463	Plastics Molding and Forming	12.5%	2	31.3%	5	16	31
428	Rubber Manufacturing	11.8%	2	11.8%	2	17	43
N/A	Miscellaneous Foods and Beverages	10.7%	3	0.0%	0	28	82
406	Grain Mills	9.1%	1	63.6%	7	11	26
410	Textile Mills <sup>a</sup>	8.3%	1	50.0%	6	12	31
411	Cement Manufacturing	4.3%	1	4.3%	1	23	48
422	Phosphate Manufacturing	0.0%	0	12.5%	1	8	14
N/A	Printing and Publishing	0.0%	0	0.0%	0	1	2

Source: ERG, 2020a.

N/A: Not Applicable

a – EPA is conducting other efforts on these categories, and they were not further reviewed in this content.

Note: Top PSCs identified through the Top Five Approach are highlighted in blue, and top PSCs identified through the Top 25 Percent Approach are highlighted in green. PSCs not included in this review due to only one facility reporting a pollutant: Coil Coating, Gum and Wood Chemicals Manufacturing, Industrial Laundries, Leather Tanning and Finishing, Tobacco Products, Carbon Black Manufacturing, Ink Formulating, Asbestos Manufacturing.

## 4. References

---

1. ERG. 2020a. Eastern Research Group, Inc. Cross-Category Concentration Analysis 2019 DMR Data Processing Documentation. (December). EPA-HQ-OW-2021-0547. DCN 11092.
2. ERG. 2020b. Eastern Research Group, Inc. 2019 DMR Industrial Monthly Average Concentration Data. (September). EPA-HQ-OW-2021-0547. DCN 11093.
3. ERG. 2020c. Eastern Research Group, Inc. 2019 DMR Industrial Monthly Average Quantity Data. (September). EPA-HQ-OW-2021-0547. DCN 11094.
4. ERG. 2020d. Eastern Research Group, Inc. 2019 DMR Flow Data. (September). EPA-HQ-OW-2021-0547. DCN 11095.
5. ERG. 2020e. Eastern Research Group, Inc. 2019 Water Pollutant Loading Tool Pollutants for Cross-Category Concentration Analysis. (September). EPA-HQ-OW-2021-0547. DCN 11096.
6. ERG. 2020f. Eastern Research Group, Inc. 2019 DMR Data Quality Review. (October). EPA-HQ-OW-2021-0547. DCN 11097.
7. ERG. 2021. Eastern Research Group, Inc. Water Pollutant Loading Tool Search Results for All Industrial Facilities in 2021. (May). EPA-HQ-OW-2021-0547. DCN 11098.
8. U.S. EPA. 2012. Technical Users Background Document for the Discharge Monitoring Report (DMR) Pollutant Loading Tool. (January). Available online at: <https://echo.epa.gov/trends/loading-tool/water-pollution-search/results/?s=546008a2bc3d6f6e2fd0ce45cb789d7d35c77c40>. EPA-HW-OW-2014-0170-0058.
9. U.S. EPA. 2020. United States Environmental Protection Agency. *EPA's Review of Industrial Wastewater Discharge Monitoring Report (DMR) Data*. (December). EPA-HQ-OW-2018-0618-0660.
10. U.S. EPA. 2021. United States Environmental Protection Agency. *Effluent Guidelines Program Plan 14*. (January). EPA-HQ-OW-2018-0618-0658.